Geometric Conditions for Subspace-Sparse Recovery
Supplementary Material

Chong You CYOU@CIS.JHU.EDU
Center for Imaging Science, Johns Hopkins University, Baltimore, MD, 21218, USA

René Vidal RVIDAL@CIS.JHU.EDU
Center for Imaging Science, Johns Hopkins University, Baltimore, MD, 21218, USA

Appendices

A. Proof of Lemma 2

Lemma. *If the inlier dictionary $\Phi \in \mathbb{R}^{n\times M}$ has full column rank, then the set of dual points, $\mathcal{D}(\Phi)$, contains exactly 2^M points specified by $\{\Phi(\Phi^\top\Phi)^{-1}\cdot \mathbf{u}, \mathbf{u}\in U_M\}$, where $U_M := \{[u_1,\cdots,u_M], u_i = \pm 1, i = 1,\cdots,M\}$.*

Proof. It can be seen in the proof of Theorem 2 that there are possibly at most 2^M dual points in the case where Φ is of full column rank. So in order to prove the result,it is enough to show that the set $\{\Phi(\Phi^\top\Phi)^{-1}\cdot \mathbf{u}, \mathbf{u}\in U_M\}$ contains 2^M points, and each of them is a dual point.

To show that there are 2^M different points, notice that U_M has 2^M points, so we are left to show that for any $\mathbf{u}_1,\mathbf{u}_2 \in U_M$ with $\mathbf{u}_1 \neq \mathbf{u}_2$, it has $\Phi(\Phi^\top\Phi)^{-1}\mathbf{u}_1 \neq \Phi(\Phi^\top\Phi)^{-1}\mathbf{u}_2$. This can be easily established by noticing that $\text{rank}(\Phi(\Phi^\top\Phi)^{-1}) = \text{rank}(\Phi) = M$, i.e., $\Phi(\Phi^\top\Phi)^{-1}$ is also of full rank, so its null space contains only the origin.

Now we show that $\Phi(\Phi^\top\Phi)^{-1}\mathbf{u}_0$ is a dual point for any $\mathbf{u}_0 \in U_M$. Denote $\eta_0 = \Phi(\Phi^\top\Phi)^{-1}\mathbf{u}_0$. By definition, we need to show that η_0 is an extreme point of the set $\mathcal{K}^o(\pm\Phi) = \{\eta \in \mathcal{R}(\Phi) : \|\Phi^\top\eta\|_\infty \leq 1\}$. First, η_0 is in $\mathcal{K}^o(\pm\Phi)$ because $\|\Phi^\top\eta_0\|_\infty = \|\mathbf{u}_0\|_\infty = 1$. Second, suppose there are two points, $\eta_1,\eta_2 \in \mathcal{K}^o(\pm\Phi)$, such that

$$\eta_0 = (1-\lambda)\eta_1 + \lambda\eta_2 \tag{1}$$

for some $\lambda \in (0,1)$, we need to show that it must be the case that $\eta_1 = \eta_2$. Notice that the columns of $\Phi(\Phi^\top\Phi)^{-1}$ span the space $\mathcal{R}(\Phi)$ and that $\eta_1,\eta_2 \in \mathcal{K}^o(\pm\Phi) \subseteq \mathcal{R}(\Phi)$, there exists $\mathbf{x}_1,\mathbf{x}_2$ such that $\eta_i = \Phi(\Phi^\top\Phi)^{-1}\mathbf{x}_i, i = 1,2$. Then by using (1), it has

$$\Phi(\Phi^\top\Phi)^{-1}\mathbf{u}_0 = (1-\lambda)\Phi(\Phi^\top\Phi)^{-1}\mathbf{x}_1 + \lambda\Phi(\Phi^\top\Phi)^{-1}\mathbf{x}_2,$$

and by left multiplying $\Phi^\top$, we have

$$\mathbf{u}_0 = (1-\lambda)\mathbf{x}_1 + \lambda\mathbf{x}_2. \tag{2}$$

Now, consider the equation for each entry separately in (2), i.e., $[\mathbf{u}_0]_i = (1-\lambda)[\mathbf{x}_1]_i + \lambda[\mathbf{x}_2]_i$, where i indexes an entry in the vector. The left hand side, being ± 1, is a extreme point of the set $[-1,1]$, while the right hand side is the convex combination of two points in $[-1,1]$, so it necessarily has that $[\mathbf{x}_1]_i = [\mathbf{x}_2]_i$. This is true for all entries i, so $\mathbf{x}_1 = \mathbf{x}_2$, thus $\eta_1 = \eta_2$, which shows that η_0 is indeed an extreme point. $\square$

B. Proof of Theorem 14

Theorem. *Given a dictionary Π. If it has $\mu(\Pi) < \frac{1}{2M-1}$, then for any partition of Π into Φ and Ψ where Φ has M columns, it has rank$(\Phi) = M$ and that PRC and DRC hold.*

Proof. If $\mu(\Pi) < 1/(2M-1)$, then it must have $\text{rank}(\Phi) = M$, this is an established result in sparse recovery. In the following, we show that PRC holds.

We start by giving an upper bound on $R(\mathcal{K}^o(\pm\Phi))$. From Lemma 2, given any $\eta \in \mathcal{K}^o(\pm\Phi)$ where $\eta \neq 0$, it can be written as $\eta = \Phi(\Phi^\top\Phi)^{-1}\mathbf{u}$ for some $\mathbf{u} \neq 0$ with $\|\mathbf{u}\|_\infty \leq 1$. Thus,

$$\|\eta\|_2^2 = \eta^\top\eta = \mathbf{u}^\top(\Phi^\top\Phi)^{-1}\mathbf{u} \leq M\cdot\frac{\mathbf{u}^\top(\Phi^\top\Phi)^{-1}\mathbf{u}}{\mathbf{u}^\top\mathbf{u}}.$$

Denote $\lambda_{\max}(\cdot), \lambda_{\min}(\cdot)$ to be the maximum and minimum eigenvalue of a symmetric matrix, respectively. We get

$$\begin{aligned}\|\eta\|_2^2 &\leq M\cdot\max_{\mathbf{u}\neq 0}\frac{\mathbf{u}^\top(\Phi^\top\Phi)^{-1}\mathbf{u}}{\mathbf{u}^\top\mathbf{u}}\\ &= M\cdot\lambda_{\max}(\Phi^\top\Phi)^{-1} = \frac{M}{\lambda_{\min}(\Phi^\top\Phi)}.\end{aligned}$$

Notice that $\Phi^\top\Phi$ is very close to identity matrix, i.e., its diagonals are 1 and the magnitude of each off-diagonal entry is bounded above by $\mu(\Pi)$. By using Gersgorin's disc theorem, $\lambda_{\min}(\Phi^\top\Phi) \geq 1-(M-1)\mu(\Pi)$, so

$$\|\eta\|_2^2 \leq \frac{M}{1-(M-1)\mu(\Pi)}.$$

As a consequence, $R(\mathcal{K}^o(\pm\Phi)) \leq \sqrt{\frac{M}{1-(M-1)\mu(\Pi)}}$.

In the second step, we give an upper bound for the right hand side of PRC. By definition,

$$\mu(\Psi, \mathcal{R}(\Phi)) = \max_{\substack{\eta\in\mathcal{R}(\Phi),\\ \|\eta\|_2=1}} \|\Psi^\top \eta\|_\infty.$$

We thus need to bound $\|\Psi^\top \eta\|_\infty$ for any $\eta \in \mathcal{R}(\Phi)$ with $\|\eta\|_2 = 1$. Consider the optimization program

$$\mathbf{x}^* = \arg\min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{s.t. } \eta = \Phi\mathbf{x}.$$

and its dual program

$$\max_{\omega} \langle \omega, \eta \rangle \quad \text{s.t. } \|\Phi^\top \omega\|_\infty \leq 1.$$

The strong duality holds since the primal problem is feasible, and the objective of the dual is bounded by $\|\omega\|_2\|\eta\|_2 \leq R(\mathcal{K}^o(\pm\Phi))$. Consequently, it has $\|\mathbf{x}^*\|_1 \leq R(\mathcal{K}^o(\pm\Phi))$. This leads to

$$\begin{aligned}\|\Psi^\top \eta\|_\infty &= \|\Psi^\top \Phi\mathbf{x}^*\|_\infty \leq \|\Psi^\top \Phi\|_\infty \|\mathbf{x}^*\|_1 \\ &\leq \mu(\Pi) R(\mathcal{K}^o(\pm\Phi)),\end{aligned}$$

in which $\|\cdot\|_\infty$ for matrix treats the matrix as a vector.

Now we combine the results from the above two parts.

$$\begin{aligned}\mu(\Psi, \mathcal{R}(\Phi)) &\leq \mu(\Pi) R(\mathcal{K}^o(\pm\Phi)) \\ &= r(\mathcal{K}(\pm\Phi))(\mu(\Pi) R(\mathcal{K}^o(\pm\Phi))^2) \\ &\leq r(\mathcal{K}(\pm\Phi)) \frac{M}{1-(M-1)\mu(\Pi)},\end{aligned}$$

in which

$$\frac{M}{1-(M-1)\mu(\Pi)} = 1 + \frac{\mu(\Pi)(2M-1)-1}{1-(M-1)\mu} < 1,$$

thus $\mu(\Psi, \mathcal{R}(\Phi)) < r(\mathcal{K}(\pm\Phi))$, which is the PRC. $\square$

References